

Sveučilište u Zagrebu  
Filozofski fakultet  
Odsjek za psihologiju

**Analiza zadataka dviju paralelnih formi Problemnog testa**

**DIPLOMSKI RAD**

Iva Bosanac

Mentor: Doc. dr. Damir Ljubotina

Zagreb, 2006.

## **SAŽETAK**

Za potrebe razredbenog postupka za studij psihologije na Filozofskom fakultetu u Zagrebu, 1994. je Alija Kulenović izradio dvije nove intencionalno paralelne forme Problemnog testa. Verzija Pb-94 je korištena 1994. a verzija Pb-95 1995. godine. Podaci dobiveni tada su korišteni u ovome radu. A. Kulenović i M.S. Žebec proveli su deskriptivnu analizu obje nove forme Problemnog testa, te provjerili njihovu paralelnost. Nikakvih drugih podataka o dvije nove forme Problemnog testa nema. U ovome radu glavni cilj je bio provesti kompletnu analizu zadataka novih formi Problemnog testa, forme Pb-94 i forme Pb-95, te faktorsku analizu. Također se predložila skraćena verzija oba testa i provjerila tendencija mijenjanja testovnih parametara sa skraćanjem testa.

Analiza zadataka ustanovila je da se zadaci u određenoj mjeri razlikuju u dvije nove forme Problemnog testa. Indeks težine u oba testa zajedno varira od 0.0371 do 0.9971 te se smanjuje prema kraju testa. Koeficijenti diskriminativne valjanosti u oba testa nalaze se u rasponu između -0.0482 i 0.4979. Interkorelacije zadataka se kreću od -0.143 do 0.466. Faktorskom analizom se i kod jednog i kod drugog testa dobiva jedan glavni faktor, nazvan "osjetljivost na probleme". Izbacivanje čak 30 zadataka iz obje forme testa rezultiralo je povećanjem težine testa (distribucija se pomiče prema srednjim vrijednostima), osjetljivost lagano opada, pouzdanost se ne mijenja bitnije kao ni faktorska struktura i korelacije s vanjskim kriterijima (testovi znanja iz psihologije, matematike i biologije). Isto tako skraćanjem testova, dvije nove forme Problemnog testa ostaju paralelne u svim verzijama s različitim brojem zadataka.

**KLJUČNE RIJEČI:** inteligencija, problemni test, analiza zadataka, faktorska analiza

## **SUMMARY**

For the needs of the entrance exam for the Psychology Department of the University of Zagreb, in 1994 Alija Kulenović made two new parallel versions of the Problem test. Collected data are used in this work. A.Kulenović and M.S.Žebec made descriptive statistic of the two new tests, and they checked the consistency of parallelism. No other information of the two new tests has been published. The major goal of current work is

item analysis, factor analysis and attempt of reducing tests and what are the consequences on the parameters of the tests.

Item analysis showed that certain amount of items are different in each test. Item difficulty varies from 0.0371 to 0.9971. Item discrimination coefficient goes from – 0.0482 until 0.4979. Intercorrelation between items are from –0.143 to 0.466. Factor analysis resulted, on both tests, with one general factor, which we called “sensitivity for problems”. Higher difficulty in both tests resulted by excluding 30 items in them, but a somewhat lower sensibility and almost no change in reliability. Factor structure and correlation with external criterions have not been changed. By excluding some items from both tests, they showed to be still parallel.

**KEYWORDS:** intelligence, problem test, item analysis, factor analysis

|                                   |           |
|-----------------------------------|-----------|
| <b>1. UVOD .....</b>              | <b>5</b>  |
| <i>O inteligenciji .....</i>      | <i>5</i>  |
| <i>O rješavanju problema.....</i> | <i>6</i>  |
| <i>O testovima .....</i>          | <i>9</i>  |
| <i>O Problemnom testu .....</i>   | <i>10</i> |
| <b>2. CILJ ISTRAŽIVANJA .....</b> | <b>13</b> |
| <b>3. METODOLOGIJA.....</b>       | <b>14</b> |
| <b>4. OBRADA REZULTATA.....</b>   | <b>15</b> |
| <i>Analiza zadataka .....</i>     | <i>16</i> |
| <i>Faktorska analiza .....</i>    | <i>19</i> |
| <i>Skraćenje testova.....</i>     | <i>22</i> |
| <b>5. RASPRAVA.....</b>           | <b>28</b> |
| <b>6. ZAKLJUČAK .....</b>         | <b>33</b> |
| <b>7. PRILOG.....</b>             | <b>34</b> |
| <b>8. LITERATURA .....</b>        | <b>36</b> |

## 1. UVOD

### *O inteligenciji*

Pri definiranju i određenju inteligencije postoje nesuglasice izazvane različitim istraživanjima i stavovima. Može se reći da univerzalne definicije i nema, dok Šverko u *Psihologijskome rječniku* navodi: "Jedan od važnijih psihologijskih konstrukata koji nije jednoznačno definiran. Najčešće označava svojstvo uspješnog snalaženja jedinke u novim situacijama, u kojima ne pomaže stereotipno nagonско ponašanje, a niti učenjem stečene navike, vještine i znanja. Za razliku od ovih neplastičnih i reproduktivnih oblika ponašanja koja su korisna samo u nekim prilikama, inteligencija je svojstvo jedinke da pronalazi nove prilagođene reakcije u prilikama bilo koje vrste." (prema Petz, 1992).

Prilikom pokušaja definiranja i istraživanja inteligencije postoje dva opća pristupa: psihometrijski (tradicionalni) pristup i alternativni pristupi.

*Psihometrijski (tradicionalni) pristup* proučava inteligenciju kao "sposobnost (snalaženja u novim situacijama, rješavanja problema, apstraktnog mišljenja, itd.) o kojoj ovisi mogući intelektualni doseg pojedinca, odnosno razina njegovih obrazovnih i profesionalnih dostignuća." (Petz, 1992). Pojedinci se razlikuju u stupnju razvijenosti tih sposobnosti, a iste mjere tzv. testovi inteligencije. No, neslaganja postoje i unutar ovoga područja s obzirom na broj i vrstu faktora od kojih bi se inteligencija sastojala. Unatoč različitim istraživanjima uz pomoć faktorske analize, rezultati nisu jednoznačni. Postoji nekoliko teoretskih pristupa:

- G-faktorski pristup - inteligencija je jedinstvena, opća kognitivna sposobnost. Pobornici ove teorije su npr. Spearman koji je svojim istraživanjima dokazivao postojanje jednoga općeg faktora koji sudjeluje u svim intelektualnim aktivnostima.
- multifaktorski pristup - inteligencija je zajednički naziv za skup različitih i različito povezanih intelektualnih sposobnosti. Taj pristup je zastupao npr. Guilford sa svojim modelom strukture intelekta: križanjem triju dimenzija (operacija, sadržaja i produkta) dobiva se 120 nezavisnih, hipotetskih faktora kognitivnih sposobnosti.
- hijerarhijski model - vrsta kompromisnog stajališta, među čijim su zastupnicima Burt i Vernon, koji pretpostavljaju da uz postojanje općeg G-faktora postoje i više užih i širih grupnih faktora intelektualnih sposobnosti.

*Alternativni pristupi*, među čijim su zastupnicima Sternberg i Berry, nastaju kao odgovor na psihometrijske pristupe, ali ni sami nisu homogeni. Naime, umjesto promatranja inteligencije kao nepromjenjive nasljedne osobine svedene na ono što mjere testovi inteligencije, oni promiču tzv. kontekstualistički pristup, gdje se inteligenciju promatra u skladu s ekološkim i kulturnim kontekstom u kojem pojedinci djeluju. Drži se da to što mjere testovi inteligencije niti je uvijek relevantno za inteligentno ponašanje jedinke u stvarnom životu, niti je jednako za sve kulture i uvjete. U skladu s tim, smatraju da je inteligencija samo naziv za sklop osobina i dispozicija koje u različitim kombinacijama rezultiraju reakcijama uspješnog prilagođavanja.

Neslaganja ima i oko drugih pitanja vezanih uz inteligenciju, kao što je npr. omjer okolinskih i nasljednih faktora u njezinu ukupnom obliku. Osim toga, inteligencija ovisi o više varijabli, a neke su od njih: spol, dob, socioekonomski status, itd... Upravo zbog svega navedenoga teško je izvesti univerzalnu definiciju inteligencije.

Kao što je naveo Zarevski (2000): "Ono što nitko razuman ne poriče u vezi s inteligencijom jest to da se javlja kao najznačajnija, ili jedna od vrlo značajnih varijabli za niz ljudskih aktivnosti: profesionalnih, socijalnih, znanstvenih, sportskih i sl., te da ima vrlo složene veze s konativnom, emocionalnom i motivacijskom strukturom i funkcioniranjem ličnosti."

### ***O rješavanju problema***

Mnogi su pokušavali opisati proces rješavanja problema nabrojajući i opisujući njegove faze. John Dewey (prema Guilford, 1971) predložio je 1910. g. klasične korake ili faze u rješavanju problema: 1) primjećuje se problem, 2) problem se locira i definira, 3) javljaju se moguća rješenja, 4) uzimaju se u obzir posljedice, 5) jedno rješenje se prihvaća. Godine 1955. D.M. Johnson (prema Guilford, 1971) pojednostavljuje proceduru rješavanja problema, tako što reducira broj faza na 3: priprema, produkcija i procjena. Merrifield i sur. 1962. (prema Guilford, 1971) predlažu model s 5 stadija: priprema, analiza, produkcija, verifikacija i reaplikacija.

Ovi koraci u rješavanju problema mogu se usporediti s koracima pri kreativnoj produkciji, gdje razlika zapravo i nema. Tako Wallas 1926. (prema Guilford, 1971)

predlaže model s 4 stadija kreativnog mišljenja: 1) priprema, 2) inkubacija, 3) iluminacija, 4) verifikacija.

Karakterističan pristup rješavanju problema, u velikoj mjeri određen načinom primanja, interpretiranja i pohranjivanja informacija, naziva se kognitivnim stilom pojedinca. “Tako je npr. utvrđeno da neki ljudi zahvaćaju širok raspon okolnih događaja ili svojstava okoline, dok se drugi koncentriraju samo na neke aspekte događaja ili neke odabrane attribute okoline. Ili, dok neki rješavaju probleme impulzivno, reagirajući na osnovi prve pretpostavke o mogućem rješenju, drugi, čiji je kognitivni stil reflektivan, sistematski nastoje doći do najboljeg rješenja, prije nego što krenu u akciju.” (Petz, 1992)

S druge strane, rigidnost označava poteškoće u promjeni usmjerenosti mišljenja, kao i u rješavanju problema na nove, različite načine. Rigidna osoba je nesposobna mijenjati strategiju rješavanja problema, kao i sagledavati probleme iz različitih gledišta. Ponašanje takve osobe tijekom rješavanja problema je nepromjenjivo i kruto. Osobine osobe koje su povezane s rigidnošću jesu: anksioznost i neurotizam, konvencionalnost i konformizam, stereotipnost i odsutnost spontanosti, regresija ili vraćanje na ranije oblike ponašanja, niska razina aspiracije ili motiva postignuća.

Uvjeti u okolini koji koče kreativnost, a potiču rigidno ponašanje jesu:

- zadaci koji usmjeravaju pojedinca da se fiksira na određenu metodu rješavanja;
- postojanje izvora frustracije i stresora za vrijeme rješavanja problema;
- činioci koji smanjuju intrinzičnu (unutrašnju) motivaciju za rješavanje problema;
- naglašavanje negativnih posljedica mogućeg neuspjeha u rješavanju problema;
- siromašno individualno iskustvo pojedinca na području rješavanja problema;
- uvođenje nepotrebnih ograničenja, prepreka i distraktora u problemsku situaciju.

Zoran Bujas je temeljnom intelektualnom sposobnošću smatrao sposobnost uočavanja i prepoznavanja problema te njihovo rješavanje. Tri su osobine koje je Bujas pritom držao bitnima: osjetljivost za probleme, fleksibilnost mišljenja te kreativnost.

Bujas je 1966. svoj temeljni koncept “*osjetljivosti za probleme*” (prema Kulenović, 1998) obrazložio ovako: “Analiziramo li procese mišljenja kojima se služi čovjek u problemnim situacijama, vidjet ćemo da uspješno snalaženje zavisi u prvom redu od njegove sposobnosti uviđanja neravnoteže u okolnim strukturama, odnosno neravnoteže

situacije i ciljeva kojima se teži.” Bujasova temeljna pretpostavka je da inteligentniji ljudi otkrivaju neravnoteže i u situacijama u kojima ju manje inteligentni ljudi ne osjećaju, tj. skloniji su uočavanju razlika između postojeće situacije i ciljeva kojima teže.

Bitan aspekt efektivnog razmišljanja u problemnim situacijama je sposobnost promatranja iste strukture iz različitih aspekata. Bujas je tu sposobnost nazvao “*misaona fleksibilnost*”. Zarevski 2000. navodi:” U najvažnije komponente Bujasove koncepcije inteligentnih ponašanja spadaju *misaona fleksibilnost* i *kreativnost*. Fleksibilnost mišljenja shvaćena je kao sposobnost uočavanja neravnoteže u različitim situacijama te u operativnom smislu nazvana osjetljivošću za probleme. Ova njegova ideja dosljedno je realizirana u nekoliko verzija Problemnog testa, a ponešto drugačije i u nekim drugim testovima kao što su Poliprofilni test i Test kombiniranih rješenja. Polazeći od pretpostavke da svi misaoni procesi (osim, donekle u mišljenju po analogiji) imaju svojstvo kreativnosti, Bujas je ovu ideju pokušao uključiti u veći broj svojih testova kao što su Test višestruke klasifikacije, Test uravnoteženih struktura i Test približnih rješenja.”

Iako je “osjetljivost za probleme” smatrana važnijom komponentom intelektualnih sposobnosti, uglavnom je zanemarivana i olako shvaćana pri testiranju inteligencije. Do zanemarivanja dolazi tako što se postavljaju zadaci sa jednakim načinom rješavanja problema, a u uputama se obično detaljno opisuje problem koji se mora riješiti te način njegova rješavanja.

Bujas je pri izradi testova uvijek nastojao uzeti u obzir dvije stvari: teorijsku podlogu intelektualnih sposobnosti kojoj se priklonio te najadekvatnije psihometrijske postupke za mjerenje te teorije. Upravo zato je nastojao svoje testove učiniti što kompleksnijima i sveobuhvatnima približivši se testovnim materijalom što više stvarnim problemnim situacijama.

Evaluacije Bujasovih testova mogu se svrstati u tri kategorije. Prva kategorija išla je u smjeru istraživanja mjere do koje je autor uspio ostvariti svoje opće i pojedinačne postavke i namjere u svakom od testova. Evaluacija u tome aspektu je pokazala da konačne verzije testova, posebno one ponuđene za praktičnu primjenu, svojom operacionalizacijom opravdavaju autorove postavke i namjere. Drugi aspekt konstruktne valjanosti odnosi se na proučavanje povezanosti testovnih rezultata i vanjskih složenih



kriterija. Evaluacija u tome aspektu posvjedočila je da pokazatelji kriterijske valjanosti koji su korišteni dosežu barem razinu modalne vrijednosti dobivenu kod drugih testova iste namjene, ali su najčešće iznad tih vrijednosti. Treća kategorija evaluacija bavi se faktorsko-analitičkim sadržajem Bujasovih testova. Sve dostupne faktorske analize koje su uključivale Bujasove testove slažu se da oni obuhvaćaju puno faktora nižeg reda, kao i jedan široki generalni faktor (Kulenović i Krizmanić, 1991).

### *O testovima*

“...(Test se) U gotovo svim znanstvenim i primijenjenim područjima koristi se kao opći naziv za praktične postupke provjeravanja ili utvrđivanja nekog stanja ili funkcionalnih karakteristika nekog mehanizma ili organizma.” (Petz, 1992). U psihologiji, test predstavlja standardizirani postupak kojim se izaziva određena aktivnost čija se veličina mjeri i vrednuje radi lakše usporedbe individualnih rezultata i smještanja istih u skupinu rezultata. Testovi se u psihologiji mogu koristiti za dijagnostičke svrhe (određivanje veličine određene psihičke karakteristike) ili za prognozu (na osnovi rezultata u testu, predviđanje budućeg uspjeha u određenim aktivnostima). Testovi se uglavnom sastoje od određenog broja čestica, gdje se uspjeh u svakoj zbraja u ukupan krajnji rezultat. Oni nemaju nultu točku jer najmanji rezultat u nekom testu ne znači odsutnost predmeta mjerenja, već samo nemogućnost testa da izmjeri toliko malu zastupljenost iste u individui.

Postoji nekoliko podjela testova. Jedna je od njih načinjena s obzirom na prirodu zadatka: testovi maksimalnog učinka te testovi tipičnih ponašanja. Nadalje, podjela se može izvesti s obzirom na način primjene testova: individualni ili grupni testovi. Ili podjela s obzirom na relevantnost vremena: testovi brzine i testovi snage. Također, testovi se mogu podijeliti i s obzirom na predmet mjerenja: testovi sposobnosti (senzornih, psihometrijskih ili mentalnih), testovi znanja (provjera efekta učenja) te testovi ličnosti (analitički ili sintetički).

### ***O Problemnom testu***

Autori originalne verzije Problemnog testa iz 1966. g. su Z. Bujas, S. Szabo i V. Kolesarić. Pomalo izmijenjen i prilagođen test, nazvan Problemni test A-forme, završen je i tiskan 1981. godine.

Problemni test ispituje domišljatost, misaonu snalažljivost i oštromnost bez obzira na školski uspjeh. Osoba kod rješavanja testa pokazuje u kojoj mjeri shvaća različite zadatke i koliko uspješno se snalazi u novim prilikama. Dakle temeljni koncept koji se mjeri Problemnim testom jest “osjetljivost za probleme”. Generirano rješenje u testovima “osjetljivosti za probleme” projicira sposobnost pojedinca da otkrije i rješava probleme. Rješavanje zadataka temelji se na principu otkrivanja problema (neravnoteže), a kada se problem jednom otkrije, rješenje je relativno jednostavno. Kratkim uputom i raznolikošću zadataka postignut je i cilj konstrukcije Problemnog testa, što je sprečavanje analogijskog mišljenja pri rješavanju testa.

Problemni test pripada kategoriji testova sposobnosti kojima se provjerava maksimalan učinak. Primjena je skupna ili pojedinačna, dobna je granica od 16 godina navise, a odgovori se upisuju u testnu knjižicu sa zadacima. Vrijeme rada ograničeno je na 45 minuta, a primjena s uputom traje oko 50 minuta. Test sadrži 70 zadataka (koji se sastoje od verbalnih, grafičkih i numeričkih oblika) te uputu. 59 zadataka je tipa nadopunjavanja, 4 ih je tipa sređivanja zadanih elemenata te 7 tipa izbacivanja suvišnih elemenata. Krković i Kolesarić (Krković, 1970) dobili su indeks težine zadataka između  $p=0.005$  do  $p=0.95$ . Hajduka (1970, prema Krković, 1970) došao je do podatka da prosječna interkorelacija zadataka iznosi svega 0.0997. Koeficijenti diskriminativne valjanosti variraju u nešto većem rasponu od 0.07 do 0.87 (prema Krković, 1970).

“Do 1970. test je primijenjen u više navrata, na različitim populacijama i s različitim ciljevima. Sažmu li se rezultati iz prikaza Krkovića i Kolesarića (1970), glavna svojstva Problemnog testa jesu: test je težinski primjeren u obrazovnom rasponu od završnog razreda osnovne škole do srednjoškolske mature; koeficijenti pouzdanosti tipa unutarnje konzistencije, određeni različitim postupcima i u različitim okolnostima variraju u rasponu od 0.85 do 0.95; stabilnost rezultata (koeficijenti test-retest pouzdanosti) u vremenskim intervalima od 21 mjesec do 3 godine ilustriraju koeficijenti u rasponu od

0.90 do 0.96; nisu pronađene razlike u učinku učenika iz gradskih i izvangradskih škola niti između muških i ženskih ispitanika; korelacije s drugim kognitivnim testovima iznose između 0.48 za verbalnu formu Testa višestruke klasifikacije do čak 0.81 za M-seriju; konačno, ustanovljene su i značajne korelacije s nekim vanjskim kriterijima (mjere školskog i akademskog uspjeha) između 0.29 i 0.53. Neke kasnije provjere kriterijske valjanosti Problemnog testa potvrdile su izložene podatke kao i pretpostavku o njihovoj djelomičnoj podcijenjenosti prilikom procjena na selekcioniranim uzorcima ispitanika (Žužul, Lugomer i Kulenović, 1984). Podaci o konstruktnoj valjanosti Problemnog testa pokazuju da se, u faktorskom smislu, najvjerojatnije radi o složenoj mjeri gdje dominira faktor koji su Krković i Kolesarić (1973) imenovali kao *osjetljivost na probleme* (3/4 varijance) uz osjetno manje udjele “*sposobnosti misaonog manipuliranja verbalnim simbolima*” i “*sposobnosti edukacije konektiva koji povezuju elemente nekog skupa*”. “ (Kulenović i Žebec, 1998).

Godine 1981. provedena je standardizacija Problemnog testa (prema Sobota, 1998) te je napravljena pomalo izmijenjena verzija testa. Sudionici su bili učenici i učenice završnih razreda srednjih škola u Republici Hrvatskoj (N = 5700 ispitanika). Raspon rezultata je bio između 0 i 67 bodova, aritmetička sredina rezultata iznosila je  $M = 35.02$  boda, a standardna devijacija rezultata  $SD = 12.15$  bodova. Tada se ujedno pokazalo da postoje statistički značajne razlike s obzirom na spol, mjesto završavanja osnovne škole te struku. Grgin (1989) je na uzorku od  $N=94$  srednjoškolaca dobio aritmetičku sredinu  $M=41$  te standardnu devijaciju  $SD=9.13$ .

S vremenom, masovna upotreba tijekom dužega vremenskog razdoblja, dovela je do zastarijevanja normi te prenaučivosti sadržaja zadataka. Stoga je 1994. za potrebe razredbenog postupka za studij psihologije na Filozofskom fakultetu u Zagrebu, Alija Kulenović izradio dvije nove i intencionalno paralelne forme Problemnog testa (Pb-94 koji se koristio za razredbeni postupak 1994. godine, te Pb-95 koji se koristio za godinu 1995.). Nastojalo se za svaki zadatak iz A-forme Problemnog testa izraditi ekvivalent s obzirom na misaoni proces i težinu zadatka. Cilj izrade nove verzije Problemnog testa jest proizvesti formu testa koja će mjeriti navedene konstrukte barem približno isto kao i A-forma.

Rezultati provjere ekvivalentnosti dviju novih formi Problemnog testa pokazuju da testovi "imaju statistički podjednake aritmetičke sredine i varijance bruto-rezultata, jednake koeficijente pouzdanosti tipa unutarnje konzistencije te da u podjednakom stupnju koreliraju s testovima znanja iz psihologije, biologije i matematike. Distribucije bruto-rezultata u Pb-94 i Pb-95 statistički su podjednakog oblika, a isti odnos dobiven je za raspodjele pokazatelja koeficijenata interne valjanosti zadataka." (Kulenović i Žebec, 1998). Dakle, može se zaključiti da su ta dva testa paralelni testovi te da se u praksi mogu kao takvi i upotrebljavati.

Nakon određenog vremena korištenja Problemnog testa u praksi, javlja se također potreba za skraćenom verzijom jer se primjena od 50 minuta s uputom pokazala predugom.

## 2. CILJ ISTRAŽIVANJA

Glavni je cilj istraživanja provesti kompletnu analizu zadataka dviju novih verzija Problemnog testa, forme Pb-94 i forme Pb-95. Otkada su konstruirana dva nova oblika Problemnog testa, takva analiza zadataka nije nigdje objavljena.

Primjena Problemnog testa u praksi ukazala je na potrebu za kraćom verzijom. U većini slučajeva 50 minuta je previše samo za test inteligencije, jer je on obično samo jedan od više testova u bateriji testova koja se primjenjuje. Stoga će se osim analize zadataka pokušati predložiti kraća verzija Problemnog testa, tj. odredit će se zadaci koji jednako dobro funkcioniraju u obje nove verzije Problemnog testa. To će biti korisno prvenstveno radi lakše i efikasnije primjene, skraćivanja vremena potrebnog za primjenu istoga. No treba također provjeriti mijenjaju li se (i kako) izbacivanjem zadataka metrijske karakteristike testa te mogu li se one poboljšati. Provjerit će se i je li paralelnost testova ostala ista nakon skraćivanja.

Također se pojavila potreba za izradom teže i lakše verzije testa radi različitosti uzorka na kojima bi se primjenjivao. Analiza zadataka provedena u ovome istraživanju uvelike će pomoći u izradi takvih verzija Problemnog testa.

Osim toga provjerit će se i konstruktna valjanost obje verzije.

Sažimajući sve navedeno, postavili smo sljedeće probleme:

- 1) provesti analizu zadataka dviju novih formi Problemnog testa: Pb-94 i Pb-95;
- 2) provjeriti faktorsku strukturu dviju novih formi Problemnog testa: Pb-94 i Pb-95;
- 3) na osnovi analize zadataka ponuditi prijedlog za skraćivanje Problemnog testa te ujedno provjeriti kako se izbacivanjem zadataka mijenja deskriptivna statistika dviju verzija tog testa.

### 3. METODOLOGIJA

Uzorak je u ovome slučaju prigodan budući da su sudionici pristupnici razredbenom postupku za psihologiju na Filozofskom fakultetu u Zagrebu. Prva forma Problemnog testa, Pb-94, primijenjena je na prijamnom postupku 1994. god., a druga forma, Pb-95, primijenjena je 1995. god.

Uzorak za Pb-94 sastoji se od N=377 sudionika, a uzorak za Pb-95 od N=273 sudionika. U oba uzorka većinska je zastupljenost žena (82% 1994. i 85% 1995. godine), što se vidi i iz Tablice 1. Sudionici oba uzorka ujednačeni su po životnoj dobi (gotovo svi su u intervalu od 18 do 20 godina) te po školskoj naobrazbi (svi imaju završenu srednju školu). Dobivena je mala razlika rezultata u korist muških sudionika (prosječno 1.22 boda) no ta razlika nije statistički značajna, što bi značilo da test podjednako funkcionira kod oba spola. Osim toga, zastupljenost muškaraca i žena u uzorku je bitno različita, što uvelike utječe na opravdanost računanja razlika među spolovima.

**Tablica 1.** Zastupljenost muškaraca i žena u uzorku pri primjeni forme Pb94 i Pb95.

|                 | <b>frekvencija<br/>Pb94</b> | <b>postotak<br/>Pb94</b> | <b>frekvencija<br/>Pb95</b> | <b>postotak<br/>Pb95</b> |
|-----------------|-----------------------------|--------------------------|-----------------------------|--------------------------|
| <b>muškarci</b> | 68                          | 18                       | 40                          | 14.7                     |
| <b>žene</b>     | 309                         | 82                       | 233                         | 85.3                     |
| <b>ukupno</b>   | 377                         | 100                      | 273                         | 100                      |

#### 4. OBRADA REZULTATA

Podaci koji su prikupljeni tijekom ispitivanja obrađeni su osobnim računalom pomoću programskog paketa SPSS za Windows verzija 11.0.

Za dvije paralelne forme Problemnog testa, dobivena je deskriptivna statistika prikazana u Tablici 2. Vizualnim pregledom uočava se da nema velikih razlika među navedenim podacima kod oba testa. “Ova opća impresija provjerena je odgovarajućim statističkim testovima te je dobiveno da su aritmetičke sredine dviju formi testa statistički podjednake ( $t=0.088$ ,  $p>0.05$ ), a isti odnos dobiven je ispitivanjem razlika njihovih varijanci ( $F=1.043$ ,  $p>0.05$ ).” (Kulenović i Žebec, 1998).

**Tablica 2.** Deskriptivna statistika dviju formi Problemnog testa.

| test  | N   | min | max | M     | SD    | $\alpha$ | prosj $r_{ij}$ | min $r_{ij}$ | max $r_{ij}$ | prosj p | prosj $r_{iu}$ |
|-------|-----|-----|-----|-------|-------|----------|----------------|--------------|--------------|---------|----------------|
| Pb-94 | 377 | 14  | 66  | 43.93 | 8.979 | 0.8655   | 0.0805         | -0.131       | 0.466        | 0.628   | 0.304          |
| Pb-95 | 273 | 16  | 65  | 43.99 | 9.169 | 0.8659   | 0.0832         | -0.143       | 0.411        | 0.629   | 0.309          |

**LEGENDA:**

N – broj ispitanika

M – aritmetička sredina

SD – standardna devijacija

$\alpha$  – Cronbachov  $\alpha$

prosj  $r_{ij}$  – prosječna interkorelacija zadataka

prosj p – prosječni indeks lakoće zadatka

prosj  $r_{iu}$  – prosječni koeficijent diskriminativnosti zadataka

Nešto manji raspon rezultata ima Pb-95, no to je moguće objasniti manjim brojem sudionika. Prosječni indeksi lakoće zadataka su kod oba testa izjednačeni te svojom veličinom upućuju da su zadaci u prosjeku nešto lakši, no taj račun uvelike ovisi o korištenom uzorku.

Koeficijenti unutarnje konzistencije, izraženi kao Cronbachov koeficijent pouzdanosti, jesu visoki te se za dva testa razlikuju tek u četvrtoj decimali ( $\alpha_{Pb94}= 0.8655$  te  $\alpha_{Pb95}= 0.8659$ ). To upućuje da su oba testa jednako visoko pouzdana.

Prosječna interkorelacija zadataka u oba testa je slična te dosta niska, što znači da su zadaci unutar testa međusobno u prosjeku nisko povezani, te se po najnižim vrijednostima interkorelacija vidi da su neki zadaci čak u statistički značajnoj negativnoj korelaciji. Korelacije testova s tri korištena vanjska kriterija su sve statistički značajne.

**Tablica 3.** Korelacije ukupnog rezultata dvaju testova s vanjskim kriterijima (testovima znanja iz psihologije, matematike i biologije).

|                              | PSIHOLOGIJA | MATEMATIKA | BIOLOGIJA |
|------------------------------|-------------|------------|-----------|
| <b>ukupni rezultat Pb-94</b> | .443**      | .420**     | .311**    |
| <b>ukupni rezultat Pb-95</b> | .424**      | .437**     | .275**    |

\*\* - korelacija je statistički značajna na razini od 0.01

### *Analiza zadataka*

Za svaki zadatak dviju formi testa, izračunat je indeks lakoće (tj. aritmetička sredina odgovora sudionika na zadatak, što je u ovome slučaju isto jer su zadaci dihotomni) te standardna devijacija. Osim toga izračunate su i korelacije svakoga zadatka s ukupnim rezultatom (koeficijent diskriminativne valjanosti zadatka) te korelacije s 3 vanjska kriterija: test znanja iz psihologije, matematike i biologije. Uz navedene podatke, u Tablici 4 nalaze se još i korelacije zadataka s glavnim faktorom (o čemu nešto više kasnije) te podaci o procijenjenoj sličnosti zadataka. Ona je određivana na temelju veličine koeficijenta diskriminativne valjanosti zadatka, tj. nespurozne korelacije zadatka s ukupnim uratkom. Odlučili smo se za umjereno strogi kriterij na temelju kojega ćemo odrediti koji zadaci su dobri a koji loši, što će pomoći u određivanju jesu li zadaci u dvije verzije istog testa slični s obzirom na kriterijsku vrijednost (ne uspoređujemo ih apsolutno već kriterijski). Dakle u ovome slučaju odlučili smo da su dobri zadaci koji imaju koeficijent diskriminativne valjanosti iznad  $r_{iu}=0.25$ , a loši kojima se on nalazi ispod navedene vrijednosti. Dobri zadaci su označeni kao “+” a loši kao “-“. Bitno je napomenuti da se tu radi o orijentacijskom kriteriju te da je granica od  $r_{iu}=0.25$  provizorno određena. Takvo određivanje sličnosti samo s obzirom na to jesu li koeficijenti diskriminativne valjanosti iznad 0.25 nije ni blizu najboljem rješenju, no ipak, procjene su napravljene radi lakšeg uvida u sličnosti kvalitete zadataka.

Pregledom Tablice 4 primjećuje se da se test Pb-94 i test Pb-95 na razini zadataka dosta razlikuju. Po osobnoj procjeni, s obzirom na veličinu koeficijenta diskriminativne valjanosti, jednako funkcionira 49 zadataka (bez obzira jesu li dobri ili loši). Kao primjer najrazličitijih zadataka navela bih zadatak broj 43. Njegov koeficijent diskriminativne valjanosti je u testu Pb-94  $r_{iu}=0.0814$  a u testu Pb-95  $r_{iu}=0.3599$ . Slična je situacija sa zadatkom 45, čiji indeks lakoće iznosi u Pb-94  $p=0.06$  a u Pb-95  $p=0.86$ , te čija se razlika može vidjeti i iz njihovih koeficijenata diskriminativne valjanosti.



**Tablica 4.** Prikaz analize zadataka za dvije forme Problemnog testa.

| zad | p <sub>Pb-94</sub> | p <sub>Pb-95</sub> | r <sub>iu</sub> Pb-94 | r <sub>iu</sub> Pb-95 | g | h | r <sub>zf</sub> Pb-94 | r <sub>zf</sub> Pb-95 | procijenjeno funkcioniranje zadataka |       |         |
|-----|--------------------|--------------------|-----------------------|-----------------------|---|---|-----------------------|-----------------------|--------------------------------------|-------|---------|
|     |                    |                    |                       |                       |   |   |                       |                       | Pb-94                                | Pb-95 | zajedno |
| 1   | 1.00               | 1.00               | .0053                 | -.0398                |   |   | .013                  | -.050                 | -                                    | -     | =       |
| 2   | .98                | .98                | .1460                 | .2097                 |   |   | .175                  | .256                  | -                                    | -     | =       |
| 3   | .96                | .91                | .1628                 | .0849                 |   |   | .184                  | .087                  | -                                    | -     | =       |
| 4   | .66                | .78                | .3295                 | .2937                 |   |   | .374                  | .319                  | +                                    | +     | =       |
| 5   | .94                | .96                | .3114                 | .1373                 |   |   | .356                  | .173                  | +                                    | -     | x       |
| 6   | .98                | .89                | .1349                 | .1551                 |   |   | .169                  | .158                  | -                                    | -     | =       |
| 7   | .87                | .63                | .2732                 | .3797                 |   |   | .303                  | .447                  | +                                    | +     | =       |
| 8   | .97                | .89                | .1472                 | .2141                 |   |   | .173                  | .260                  | -                                    | -     | =       |
| 9   | .74                | .84                | .4013                 | .2883                 |   |   | .433                  | .344                  | +                                    | +     | =       |
| 10  | .55                | .64                | -.0482                | .1712                 |   |   | -.054                 | .216                  | -                                    | -     | =       |
| 11  | .75                | .81                | .1920                 | .1948                 |   |   | .224                  | .217                  | -                                    | -     | =       |
| 12  | .56                | .48                | .1941                 | .2138                 |   |   | .221                  | .243                  | -                                    | -     | =       |
| 13  | .34                | .51                | .2748                 | .2662                 | + |   | .315                  | .291                  | +                                    | +     | =       |
| 14  | .63                | .93                | .4330                 | .3129                 | + |   | .499                  | .356                  | +                                    | +     | =       |
| 15  | .95                | .63                | .1731                 | .1458                 |   | + | .210                  | .154                  | -                                    | -     | =       |
| 16  | .83                | .92                | .4485                 | .3536                 | + |   | .508                  | .421                  | +                                    | +     | =       |
| 17  | .89                | .86                | .1296                 | .1044                 |   |   | .154                  | .128                  | -                                    | -     | =       |
| 18  | .97                | .82                | .1649                 | .3369                 |   |   | .190                  | .387                  | -                                    | +     | x       |
| 19  | .80                | .73                | .1562                 | .1594                 |   |   | .188                  | .188                  | -                                    | -     | =       |
| 20  | .53                | .86                | .1859                 | .2009                 |   |   | .216                  | .239                  | -                                    | -     | =       |
| 21  | .43                | .55                | .2654                 | .3164                 |   |   | .293                  | .358                  | +                                    | +     | =       |
| 22  | .70                | .30                | .2101                 | .2632                 |   |   | .226                  | .301                  | -                                    | +     | x       |
| 23  | .81                | .90                | .1810                 | .3005                 |   |   | .220                  | .349                  | -                                    | +     | x       |
| 24  | .68                | .77                | .4419                 | .4424                 | + |   | .499                  | .495                  | +                                    | +     | =       |
| 25  | .65                | .52                | .4009                 | .2392                 | + |   | .446                  | .278                  | +                                    | -     | x       |
| 26  | .84                | .66                | .2668                 | .2716                 |   |   | .313                  | .304                  | +                                    | +     | =       |
| 27  | .92                | .81                | .3174                 | .4249                 | + |   | .365                  | .491                  | +                                    | +     | =       |
| 28  | .90                | .83                | .2497                 | .3702                 |   | + | .290                  | .436                  | -                                    | +     | x       |
| 29  | .17                | .46                | -.0197                | .1866                 |   |   | -.008                 | .213                  | -                                    | -     | =       |
| 30  | .67                | .78                | .2445                 | .2224                 |   |   | .295                  | .265                  | -                                    | -     | =       |
| 31  | .55                | .66                | .3225                 | .1435                 | + |   | .381                  | .158                  | +                                    | -     | x       |
| 32  | .77                | .48                | .2007                 | .2128                 |   |   | .232                  | .234                  | -                                    | -     | =       |
| 33  | .85                | .81                | .2406                 | .1656                 |   |   | .283                  | .192                  | -                                    | -     | =       |
| 34  | .88                | .82                | .1988                 | .3026                 |   | + | .246                  | .360                  | -                                    | +     | x       |
| 35  | .80                | .90                | .1051                 | .2698                 |   |   | .122                  | .278                  | -                                    | +     | x       |
| 36  | .83                | .61                | .2843                 | .1551                 |   |   | .322                  | .179                  | +                                    | -     | x       |
| 37  | .65                | .65                | .4937                 | .4430                 | + |   | .552                  | .501                  | +                                    | +     | =       |
| 38  | .69                | .72                | .3264                 | .3036                 |   |   | .373                  | .357                  | +                                    | +     | =       |
| 39  | .80                | .83                | .2646                 | .2653                 |   |   | .312                  | .311                  | +                                    | +     | =       |
| 40  | .81                | .98                | .2205                 | .1068                 |   |   | .249                  | .150                  | -                                    | -     | =       |
| 41  | .63                | .45                | .3521                 | .3407                 |   |   | .295                  | .393                  | +                                    | +     | =       |
| 42  | .62                | .56                | .3077                 | .3506                 |   |   | .342                  | .410                  | +                                    | +     | =       |
| 43  | .18                | .39                | .0814                 | .3599                 |   |   | .104                  | .413                  | -                                    | +     | x       |
| 44  | .60                | .71                | .4979                 | .4793                 | + |   | .564                  | .536                  | +                                    | +     | =       |
| 45  | .06                | .86                | .0793                 | .2738                 |   |   | .086                  | .319                  | -                                    | +     | x       |
| 46  | .23                | .37                | .1208                 | .2460                 |   |   | .132                  | .276                  | -                                    | -     | =       |
| 47  | .62                | .77                | .2598                 | .2539                 |   |   | .299                  | .294                  | +                                    | +     | =       |

|    |     |     |       |       |   |   |      |      |   |   |   |
|----|-----|-----|-------|-------|---|---|------|------|---|---|---|
| 48 | .84 | .73 | .3110 | .4418 | + |   | .357 | .510 | + | + | = |
| 49 | .78 | .91 | .2044 | .1119 |   |   | .230 | .141 | - | - | = |
| 50 | .55 | .33 | .3643 | .2170 | + |   | .426 | .247 | + | - | x |
| 51 | .63 | .77 | .4493 | .4694 | + | + | .515 | .539 | + | + | = |
| 52 | .48 | .45 | .2657 | .0393 |   |   | .311 | .034 | + | - | x |
| 53 | .81 | .74 | .2813 | .2765 |   |   | .330 | .336 | + | + | = |
| 54 | .56 | .61 | .4613 | .3520 | + |   | .520 | .391 | + | + | = |
| 55 | .36 | .37 | .3516 | .2775 | + |   | .396 | .321 | + | + | = |
| 56 | .55 | .32 | .4059 | .3273 | + |   | .459 | .353 | + | + | = |
| 57 | .07 | .18 | .2703 | .3049 |   |   | .299 | .335 | + | + | = |
| 58 | .46 | .32 | .4215 | .3860 | + |   | .476 | .427 | + | + | = |
| 59 | .54 | .25 | .3893 | .2491 |   |   | .440 | .282 | + | - | x |
| 60 | .54 | .37 | .3857 | .2938 |   |   | .431 | .331 | + | + | = |
| 61 | .42 | .36 | .3887 | .2431 | + |   | .441 | .282 | + | - | x |
| 62 | .63 | .66 | .2496 | .3187 |   |   | .277 | .348 | - | + | x |
| 63 | .84 | .72 | .2020 | .3284 |   | + | .246 | .396 | - | + | x |
| 64 | .14 | .17 | .2329 | .3058 |   |   | .256 | .342 | - | + | x |
| 65 | .32 | .23 | .2161 | .2817 |   |   | .244 | .310 | - | + | x |
| 66 | .04 | .12 | .2178 | .1636 |   |   | .249 | .195 | - | - | = |
| 67 | .22 | .30 | .3785 | .4018 |   |   | .423 | .449 | + | + | = |
| 68 | .37 | .29 | .2417 | .3164 |   |   | .275 | .356 | - | + | x |
| 69 | .36 | .41 | .4541 | .4587 | + | + | .504 | .517 | + | + | = |
| 70 | .20 | .18 | .3484 | .3433 |   | + | .396 | .383 | + | + | = |

**LEGENDA:**

p – indeks lakoće

$r_{iu}$  – korigirana korelacija zadatka s ukupnim rezultatom (koef. diskriminativne valjanosti)

g – zadaci najbolje povezani sa sva tri vanjska kriterija (s razinom rizika od 0.01) za Pb94

h – zadaci najbolje povezani sa sva tri vanjska kriterija (s razinom rizika od 0.01) za Pb95

$r_{zf}$  – povezanost zadataka s 1.faktorom prije rotacije

+ dobri zadaci

- loši zadaci

= zadaci jednako funkcioniraju u oba testa s obzirom na zadani kriterij

x zadaci različito funkcioniraju u oba testa s obzirom na zadani kriterij

Najteži zadatak u oba testa je broj 66 ( $p_{Pb-94}=0.0371$ ,  $p_{Pb-95}=0.1209$ ), a najlakši je prvi ( $p_{Pb-94}=0.9971$ ,  $p_{Pb-95}=0.9963$ ). Općenito se indeks lakoće zadataka smanjuje od početka ka kraju, što znači da je test teži prema kraju iako ima onih koji su iznimka (npr. 13. zadatak je nešto teži za prvi dio testa, a onaj broj 63 je lagan s obzirom na ostale zadatke u zadnjem dijelu testa). Kao što je već navedeno, kod određenih zadataka se indeks lakoće znatno razlikuje u dva testa. Neki od njih su: 22, 29, 32, 43, 59, no ekstrem bi bio već navedeni pod brojem 45.

Koeficijenti diskriminativne valjanosti se također u velikom broju zadataka razlikuju. Za test Pb-94, najmanja vrijednost koeficijenta je za 10. zadatak ( $r_{iu}=-0.0482$ ) a najmanja apsolutna vrijednost je za 1. zadatak ( $r_{iu}=0.0053$ ), dok kod testa Pb-95 najmanji koeficijent ima 1. zadatak ( $r_{iu}=-0.0398$ ) a najmanju apsolutnu vrijednost ima 52. zadatak

( $r_{iu}=0.0393$ ). Najviše koeficijente diskriminativne valjanosti ima 44. zadatak u oba testa (za Pb-94  $r_{iu}=0.4979$ , za Pb-95  $r_{iu}=0.4793$ ). Upravo on je primjer najbolje postavljenog zadatka u oba testa, tj. onoga koji najbolje razlikuje uspješne od neuspješnih.

Prosječna interkorelacija zadataka kod Pb-94 iznosi  $r_{ij}=0.07317$  a kod Pb-95  $r_{ij}=0.08133$ . U testu Pb-94 vrijednosti interkorelacija kreću se od  $r_{ij}=-0.131$  do  $r_{ij}=0.466$ . Kod testa Pb-95 one se kreću od  $r_{ij}=-0.143$  do  $r_{ij}=0.411$ .

Zadataka koji su sa sva 3 vanjska kriterija statistički značajno povezani, u testu Pb-94 ima 18, a u testu Pb-95 svega 7. U oba testa samo su zadaci 51 i 69 statistički značajno povezani sa sva tri vanjska kriterija. Moguće je da se statistička povezanost zadatka sa sva tri vanjska kriterija pojavi i po slučaju, no ako je isti zadatak u oba testa značajan u tome aspektu, to već znači da je zadatak visoko valjan u oba testa te u kontekstu samog konstrukta.

### ***Faktorska analiza***

Da bi se provjerila postojanost povezanosti skupa varijabli (zadataka) te etiologija njihove eventualne povezanosti, primjerena obrada rezultata je faktorska analiza. Njezin glavni cilj je utvrđivanje postojanja zajedničkih izvora kovariranja (latentnih varijabli, faktora) pomoću kojih se želi objasniti što veći dio kovarijabiliteta manifestnih varijabli te odrediti stupanj povezanosti svake korištene varijable s dobivenim faktorima. Dakle, cilj nam je provjeriti faktorsku strukturu unutar testa, tj. faktorsku valjanost. Kao osnova za faktorsku analizu poslužile su međusobne korelacije 70 zadataka.

Prije uporabe faktorske analize potrebno je provjeriti pogodnost korelacijske matrice za provedbu faktorske analize. Postoji nekoliko standardiziranih postupaka kojima se to radi. Mi ćemo koristiti dva postupka: Bartlettov test te Kaiser-Meyer-Olkinov test.

**Tablica 5.** KMO i Bartlettov test za dvije forme Problemnog testa.

|                                  | <b>Pb-94</b> | <b>Pb-95</b> |
|----------------------------------|--------------|--------------|
| <b>Kaiser-Meyer-Olkin indeks</b> | 0.747        | 0.707        |
| <b>Bartlettov test</b>           | 4729.372**   | 3986.073**   |

\*\*statistički značajno na razini 0.01

Bartlettov test za oba testa je statistički značajan što znači da su i korelacijske matrice statistički značajne, tj. da se dobivena korelacijska matrica razlikuje od matrice identiteta. Takav rezultat Bartlettovog testa mogao bi biti uzrokovan velikim brojem ispitanika, što može pokazati da je matrica pogodna za faktorizaciju iako ona to ne mora biti. Stoga je izračunat i KMO indeks, koji je za oba testa dobar te govori da je korelacijska matrica prihvatljiva za faktorsku analizu, iako veličina tog indeksa i nije osobito visoka.

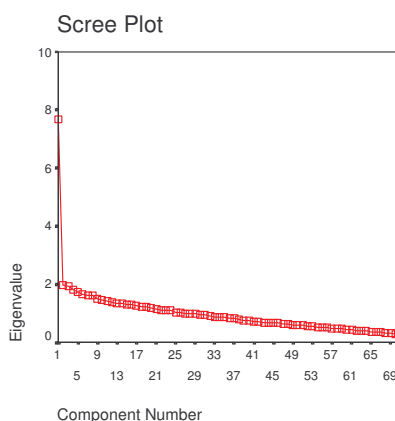
**Tablica 6.** Ekstrahirani faktori i njihovi karakteristični korijeni pod modelom glavnih komponenata za oba testa.

| <b>komponenta</b> | <b>karakteristični korijeni za Pb-94</b> | <b>% ukupne varijance objašnjeno pojedinim faktorom za Pb-94</b> | <b>karakteristični korijeni za Pb-95</b> | <b>% ukupne varijance objašnjeno pojedinim faktorom za Pb-95</b> |
|-------------------|--|--|--|--|
| 1                 | 7.659                                    | 10.941   | 7.602                                    | 10.860   |
| 2                 | 1.986                                    | 2.837  | 2.260                                    | 3.229  |
| 3                 | 1.937                                    | 2.767  | 2.093                                    | 2.989  |
| 4                 | 1.811                                    | 2.588  | 1.891                                    | 2.702  |
| 5                 | 1.743                                    | 2.491  | 1.846                                    | 2.637  |
| 6                 | 1.672                                    | 2.389  | 1.788                                    | 2.554  |
| 7                 | 1.630                                    | 2.328  | 1.769                                    | 2.527  |
| 8                 | 1.602                                    | 2.288  | 1.672                                    | 2.389  |
| 9                 | 1.501                                    | 2.145  | 1.622                                    | 2.317  |
| 10                | 1.458                                    | 2.082  | 1.603                                    | 2.290  |
| 11                | 1.437                                    | 2.053  | 1.562                                    | 2.232  |
| 12                | 1.371                                    | 1.959  | 1.486                                    | 2.123  |
| 13                | 1.340                                    | 1.915  | 1.423                                    | 2.033  |
| 14                | 1.327                                    | 1.896  | 1.381                                    | 1.973  |
| 15                | 1.301                                    | 1.859  | 1.351                                    | 1.930  |
| 16                | 1.280                                    | 1.829  | 1.323                                    | 1.889  |
| 17                | 1.260                                    | 1.800  | 1.282                                    | 1.831  |
| 18                | 1.221                                    | 1.744  | 1.253                                    | 1.789  |
| 19                | 1.216                                    | 1.737  | 1.232                                    | 1.759  |
| 20                | 1.185                                    | 1.693  | 1.209                                    | 1.727  |
| 21                | 1.147                                    | 1.639  | 1.142                                    | 1.632  |
| 22                | 1.118                                    | 1.598  | 1.121                                    | 1.602  |
| 23                | 1.108                                    | 1.582  | 1.117                                    | 1.596  |
| 24                | 1.095                                    | 1.564  | 1.080                                    | 1.543  |
| 25                | 1.036                                    | 1.481  | 1.069                                    | 1.527  |
| 26                | 1.024                                    | 1.462  | 1.054                                    | 1.505  |
| 27                |  |  | 1.031                                    | 1.473  |

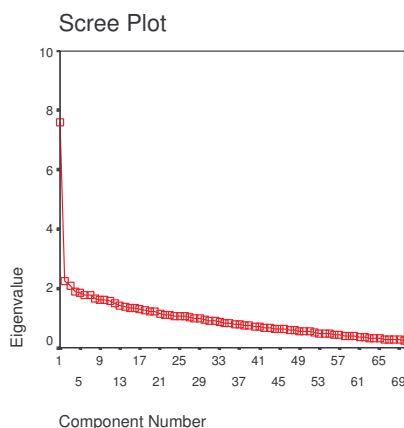
Provedena je faktorska analiza oba testa pod modelom glavnih komponenata, te su dobiveni rezultati prikazani u Tablici 6. Uz Guttman-Kaiserov kriterij, u testu Pb-94 dobiveno je 26 faktora, a u testu Pb-95 27 faktora. No, veličine njihovih karakterističnih korjenova ukazuju na jedan faktor, pošto se ostalima ne razlikuju značajno vrijednosti

korjenova. To se najbolje može primijetiti iz grafičkih prikaza 1 i 2, gdje su prikazani scree-plotovi karakterističnih korjenova faktora u svakom od testa. Rotacija nije provedena.

**Graf 1.** Grafički prikaz vrijednosti karakterističnih korjenova za faktore u testu Pb-94.



**Graf 2.** Grafički prikaz vrijednosti karakterističnih korjenova za faktore u testu Pb-95.



Iako prvi faktor u oba testa ne objašnjava veliki dio ukupne varijance (10.941% u Pb-94 te 10.860% u Pb-95), on jedini vizualno i brojčano iskače među svim ostalim faktorima u oba testa. Zasićenja zadataka u oba testa s glavnim faktorom nalaze se u Tablici 4.

Za kraj bi se moralo navesti da se u Problemnome testu radi o binarnim zadacima, što znači da je varijanca u zadacima smanjena te da su njihove međusobne korelacije niže nego da su zadaci na nekoj metrički superiornijoj skali. To dovodi do hiperfaktORIZACIJE, tj. do pojavljivanja većeg broja faktora s karakterističnim korijenom iznad 1.

### ***Skraćenje testova***

Kao temelj odabira zadataka koje bi se trebalo izbaciti prvenstveno služe koeficijenti diskriminativne valjanosti. Uostalom, definicija tog koeficijenta glasi da on prikazuje koliko je određeni zadatak valjan da razlikuje uspješne od neuspješnih pojedinaca. Visoki koeficijent znači da zadatak izvrsno diskriminira sudionike s obzirom na njihovu sposobnost koja se mjeri testom. Kao još jedan pokazatelj valjanosti zadatka može poslužiti i zasićenje glavnim faktorom. Ono ukazuje koliko je određeni zadatak povezan s glavnim faktorom koji se pokazao značajnim u korištenim testovima. Izbacivanjem zadataka koji nisu povezani s njim test se faktorski “čisti”. Ujedno su od velike koristi podaci o promjeni u Cronbachovom koeficijentu pouzdanosti ako se određeni zadatak izbací. Tablica s tim podacima nalazi se u prilogu.

U Tablici 4 dane su procjene funkcioniranja pojedinih zadataka zasnovane na koeficijentima diskriminativne valjanosti, koje će najviše poslužiti za odabir zadataka u skraćenim verzijama.

U nastavku pokušalo se u nekoliko koraka postupno u oba testa izbacivati skupine od po 5 zadataka te provjeravati kako se izbacivanjem mijenjaju parametri u oba testa. Dobiveni rezultati nalaze se u Tablici 8 i Tablici 9. Zadatak se izbacuje ako podjednako loše funkcionira u oba testa ili loše u jednom, a osrednje u drugom. Na kraju, izbacuju se zadaci s lošim funkcioniranjem u jednom testu, a dobrim u drugom. Raspored izbacivanja zadataka nalazi se u Tablici 7.

**Tablica 7.** Raspored izbacivanja zadataka.

| verzija testa (broj zadataka) | redni broj izbačenih zadataka | procijenjeno vrijeme primjene bez upute |
|-------------------------------|-------------------------------|---|
| 70 zadataka                   |                               | 45min                                   |
| 65 zadataka                   | 1, 3, 10, 17, 29              | 42 min                                  |
| 60 zadataka                   | 6, 11, 15, 19, 49             | 39 min                                  |
| 55 zadataka                   | 2, 8, 20, 40, 66              | 36 min                                  |
| 50 zadataka                   | 12, 32, 45, 46, 52            | 33 min                                  |
| 45 zadataka                   | 22, 33, 35, 36, 65            | 30 min                                  |
| 40 zadataka                   | 5, 18, 23, 31, 34             | 27 min                                  |

U Tablici 7 se ujedno nalaze i procjene vremena potrebnog za rješavanje svake od verzija testova. U idućih nekoliko tablica prikazana je deskriptivna statistika novih, skraćenih verzija Problemnog testa za svaki od dva paralelna oblika tog testa.

**Tablica 8.** Deskriptivna statistika ukupnog rezultata testa Pb-94 s različitim brojem zadataka.

| verzija | M      | SD     | $\alpha$ | prosj $r_{ij}$ | prosj p | r s psiho | r s mate | r s bio |
|---------|--------|--------|----------|----------------|---------|-----------|----------|---------|
| 70 zad  | 43.93  | 8.979  | .8655    | 0.0732         | 0.628   | .443      | .420     | .311    |
| 65 zad  | 40.363 | 8.9054 | .8707    | .0912          | .6210   | .450      | .420     | .312    |
| 60 zad  | 36.109 | 8.5783 | .8688    | .0971          | .6018   | .451      | .417     | .309    |
| 55 zad  | 32.772 | 8.2830 | .8665    | .1035          | .5959   | .446      | .414     | .308    |
| 50 zad  | 30.679 | 7.8443 | .8657    | .1126          | .6136   | .456      | .414     | .312    |
| 45 zad  | 27.186 | 7.3851 | .8639    | .1218          | .6041   | .462      | .421     | .309    |
| 40 zad  | 23.037 | 6.9319 | .8577    | .1297          | .5759   | .459      | .422     | .307    |

**LEGENDA:**

M – aritmetička sredina

SD – standardna devijacija

$\alpha$  – Cronbachov koeficijent pouzdanosti

prosj  $r_{ij}$  – prosječna interkorelacija zadataka

prosj p – prosječni indeks lakoće

r s psiho – korelacija s ispitom znanja iz psihologije

r s mate – korelacija s ispitom znanja iz matematike

r s bio – korelacija s ispitom znanja iz biologije

**Tablica 9.** Deskriptivna statistika ukupnog rezultata testa Pb-95 s različitim brojem zadataka.

| verzija | M      | SD     | $\alpha$ | prosj $r_{ij}$ | prosj p | r s psiho | r s mate | r s bio |
|---------|--------|--------|----------|----------------|---------|-----------|----------|---------|
| 70 zad  | 43.989 | 9.169  | .8659    | 0.0813         | 0.629   | .424      | .437     | .275    |
| 65 zad  | 40.125 | 8.8968 | .8665    | .0915          | .6173   | .425      | .443     | .280    |
| 60 zad  | 36.154 | 8.5486 | .8658    | .0984          | .6026   | .420      | .442     | .272    |
| 55 zad  | 32.326 | 8.2924 | .8631    | .1050          | .5877   | .414      | .440     | .259    |
| 50 zad  | 29.678 | 7.7814 | .8622    | .1124          | .5936   | .411      | .445     | .263    |
| 45 zad  | 26.821 | 7.2898 | .8587    | .1205          | .5960   | .414      | .425     | .268    |
| 40 zad  | 22.649 | 6.8219 | .8520    | .1279          | .5662   | .411      | .427     | .270    |

**LEGENDA:**

M – aritmetička sredina

SD – standardna devijacija

$\alpha$  – Cronbachov koeficijent pouzdanosti

prosj  $r_{ij}$  – prosječna interkorelacija zadataka

prosj p – prosječni indeks lakoće

r s psiho – korelacija s ispitom znanja iz psihologije

r s mate – korelacija s ispitom znanja iz matematike

r s bio – korelacija s ispitom znanja iz biologije

Aritmetička sredina oba testa se sa smanjenjem broja zadataka smanjuje, što je bilo za očekivati, no veličina smanjenja je povezana i sa smanjenjem prosječnog indeksa lakoće

zadataka. Cronbachov  $\alpha$  u testu Pb-94 raste do verzije od 50 zadataka, što bi značilo da izbacivanjem čak 20 zadataka koeficijent pouzdanosti raste. U testu Pb-95 koeficijent pouzdanosti je veći samo u verziji od 65 zadataka, dok u ostalim verzijama opada. No razlike u pouzdanosti testova svih duljina nisu velike. Prosječna interkorelacija zadataka s izbacivanjem zadataka konstantno raste u oba testa, a korelacije oba testa s tri vanjska kriterija se uz male pomake uglavnom održavaju stalnima. U tablicama 10 i 11 se nalaze podaci vezani uz osjetljivost.

**Tablica 10.** Podaci vezani uz osjetljivost skraćenih verzija testa Pb-94.

| verzija | min | max | raspon | BRR | BOR   | $\delta$ |
|---------|-----|-----|--------|-----|-------|----------|
| 70 zad  | 14  | 66  | 52     | 47  | 68683 | 0.98     |
| 65 zad  | 11  | 61  | 50     | 44  | 68613 | 0.98     |
| 60 zad  | 9   | 56  | 47     | 44  | 68582 | 0.981    |
| 55 zad  | 7   | 52  | 45     | 44  | 68506 | 0.982    |
| 50 zad  | 6   | 48  | 42     | 41  | 68356 | 0.981    |
| 45 zad  | 6   | 44  | 38     | 36  | 68270 | 0.982    |
| 40 zad  | 2   | 39  | 37     | 36  | 68130 | 0.983    |

**LEGENDA:**

min – najmanji dobiveni rezultat

max – maksimalni dobiveni rezultat

BRR – broj različitih rezultata

BOR – broj ostvarenih razlikovanja

$\delta$  – Fergusonov koeficijent osjetljivosti

**Tablica11.** Podaci vezani uz osjetljivost skraćenih verzija testa Pb-95.

| verzija | min | max | raspon | BRR | BOR   | $\delta$ |
|---------|-----|-----|--------|-----|-------|----------|
| 70 zad  | 16  | 65  | 49     | 45  | 35983 | 0.979    |
| 65 zad  | 15  | 60  | 45     | 42  | 35970 | 0.98     |
| 60 zad  | 12  | 55  | 43     | 41  | 35871 | 0.979    |
| 55 zad  | 9   | 51  | 42     | 41  | 35917 | 0.981    |
| 50 zad  | 9   | 47  | 38     | 38  | 35807 | 0.98     |
| 45 zad  | 8   | 43  | 35     | 34  | 35688 | 0.979    |
| 40 zad  | 5   | 38  | 33     | 34  | 35584 | 0.979    |

**LEGENDA:**

min – najmanji dobiveni rezultat

max – maksimalni dobiveni rezultat

BRR – broj različitih rezultata

BOR – broj ostvarenih razlikovanja

$\delta$  – Fergusonov koeficijent osjetljivosti

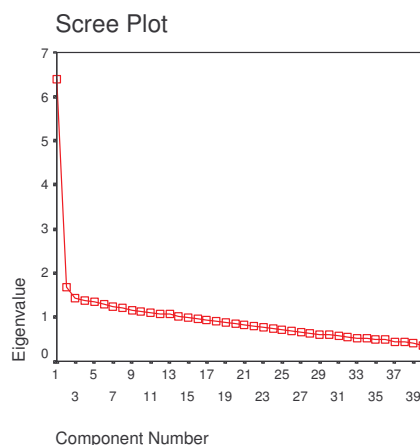
Raspon rezultata smanjuje se skraćanjem testa, te je ono postupno i gotovo simetrično u oba testa. Vrijednosti ukupnih rezultata općenito idu prema nižim razinama. Standardna devijacija rezultata također se smanjuje. Oblik distribucije ne mijenja se bitno. Da bi se provjerilo razlikuje li se distribucija rezultata za broj zadataka 40 od normalne



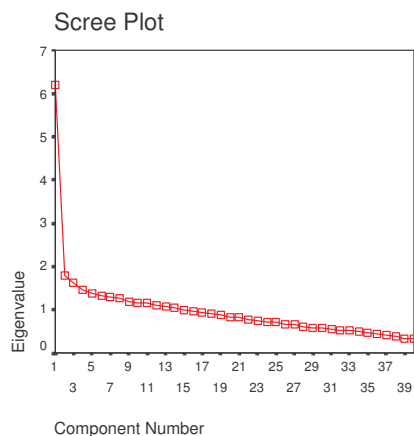
distribucije, korišten je Kolmogorov-Smirnovljev test. On je pokazao da se distribucije rezultata testova Pb-94 i Pb-95 od 40 zadataka statistički značajno ne razlikuju od normalne distribucije, što je slučaj i kod verzija od 70 zadataka (Pb-94, 40 zad.: max  $d=0.053$ ,  $p<0.05$ ; Pb-95, 40 zad: max  $d=0.081$ ,  $p<0.05$ ). BOR se smanjuje, što je i očekivano jer se smanjuje broj zadataka. S druge strane, Fergusonov koeficijent osjetljivosti uopće se ne mijenja te se čak u testu Pb-94 primjećuje njegov slabi rast.

Faktorska struktura se promjenom broja zadataka bitno ne mijenja. Uglavnom dolazi do smanjenja ukupnog broja faktora kojima je karakteristični korijen veći od 1, no samo jedan faktor se i dalje ističe. U testu Pb-94 nešto više se počinje isticati drugi faktor, no njegov odmak od ostalih faktora nije toliko značajan. Na Grafu 3 i Grafu 4 prikazani su scree-plotovi faktorskih rješenja za oba testa u slučaju sa 40 zadataka.

**Graf 3.** Grafički prikaz vrijednosti karakterističnih korjenova za faktore u testu Pb-94 od 40 zadataka.



**Graf 4.** Grafički prikaz vrijednosti karakterističnih korjenova za faktore u testu Pb-95 od 40 zadataka.



Nakon izbacivanja određenih zadataka potrebno je provjeriti jesu li skraćene verzije testova i dalje ostale paralelne. U ovome radu pokušali smo provjeriti podjednakost određenih parametara za svaku verziju skraćivanja testova, što je zadano samom teorijom o paralelnosti testova. Dakle, provjerit ćemo podjednakost: aritmetičkih sredina, varijanci bruto-rezultata te korelacija s vanjskim kriterijskim varijablama. Dodatno bi se trebala provjeriti podjednakost međusobnih korelacija (koeficijenta pouzdanosti), no budući da se tu radi o dva nezavisna skupa podataka to nije moguće direktno provjeriti. Umjesto toga su uspoređivani koeficijenti pouzdanosti dvaju testova, Cronbachovog  $\alpha$ , jer ako su dva testa paralelna, onda bi njihovi koeficijenti pouzdanosti tipa unutarnje konzistencije trebali biti podjednaki.

**Tablica 10.** Usporedni prikaz određenih parametara za svaku od skraćenih verzija testova Pb-94 i Pb-95.

| verzija | aritmetička sredina |      | varijanca |      | Cronbach $\alpha$ |       | r s psiho |        | r s mate |        | r s bio |        |
|---------|---------------------|------|-----------|------|-------------------|-------|-----------|--------|----------|--------|---------|--------|
|         | t                   | p    | F         | p    | Pb-94             | Pb-95 | Pb-94     | Pb-95  | Pb-94    | Pb-95  | Pb-94   | Pb-95  |
| 70 zad  | -.088               | .930 | .007      | .933 | .8655             | .8659 | .443**    | .424** | .420**   | .437** | .311**  | .275** |
| 65 zad  | .338                | .736 | .044      | .833 | .8707             | .8665 | .450**    | .425** | .420**   | .443** | .312**  | .280** |
| 60 zad  | -.066               | .947 | .073      | .787 | .8688             | .8658 | .451**    | .420** | .417**   | .442** | .309**  | .272** |
| 55 zad  | .677                | .499 | .046      | .829 | .8665             | .8631 | .446**    | .414** | .414**   | .440** | .308**  | .259** |
| 50 zad  | 1.612               | .107 | .108      | .743 | .8657             | .8622 | .456**    | .411** | .414**   | .445** | .312**  | .263** |
| 45 zad  | .626                | .532 | .184      | .668 | .8639             | .8587 | .462**    | .414** | .421**   | .425** | .309**  | .268** |
| 40 zad  | .710                | .478 | .286      | .593 | .8577             | .8520 | .459**    | .411** | .422**   | .427** | .307**  | .270** |

**LEGENDA:**

\*\*statistički značajno na razini 0.05

t – vrijednost dobivena t-testom za dva nezavisna uzorka

p – stupanj značajnosti razlike

F – vrijednost dobivena Levenovim testom razlike među varijancama

r s psiho – korelacija s ispitom znanja iz psihologije

r s mate – korelacija s ispitom znanja iz matematike

r s bio – korelacija s ispitom znanja iz biologije

U svim se skraćenim oblicima testova Pb-94 i Pb-95 aritmetičke sredine i varijance statistički značajno ne razlikuju. Isto tako se može pretpostaviti da se Cronbachovi koeficijenti pouzdanosti te korelacije sa sva tri vanjska kriterija znatnije ne razlikuju između dva testa. Za vanjske kriterije (testove znanja) bitno je navesti da su oni korišteni 1994. i 1995. godine donekle različitih psihometrijskih svojstava (Žebec, 1997).

Dodatno smo izračunali korelacije između svih verzija testova za svaki od dva korištena testa, jer je pri primjeni kraćih verzija korisno vidjeti koliko ta verzija korelira s originalnom, cijelom verzijom. U nastavku su tablice 11 i 12, gdje se nalaze korelacijske matrice svih verzija svakog od testa. Kao što se vidi, sve korelacije su izrazito visoke.

**Tablica 11.** Korelacijska matrica svih verzija testa Pb-94.

| verzije testa | 70 zad | 65 zad | 60 zad | 55 zad | 50 zad | 45 zad | 40 zad |
|---------------|--------|--------|--------|--------|--------|--------|--------|
| 70 zad        | 1      | .997** | .994** | .991** | .986** | .976** | .973** |
| 65 zad        |        | 1      | .997** | .994** | .988** | .979** | .975** |
| 60 zad        |        |        | 1      | .997** | .991** | .984** | .979** |
| 55 zad        |        |        |        | 1      | .994** | .986** | .982** |
| 50 zad        |        |        |        |        | 1      | .993** | .988** |
| 45 zad        |        |        |        |        |        | 1      | .995** |
| 40 zad        |        |        |        |        |        |        | 1      |

\*\* statistički značajno na razini 0.01

**Tablica 12.** Korelacijska matrica svih verzija testa Pb-95.

| verzije testa | 70 zad | 65 zad | 60 zad | 55 zad | 50 zad | 45 zad | 40 zad |
|---------------|--------|--------|--------|--------|--------|--------|--------|
| 70 zad        | 1      | .996** | .992** | .990** | .983** | .974** | .968** |
| 65 zad        |        | 1      | .995** | .992** | .984** | .974** | .969** |
| 60 zad        |        |        | 1      | .998** | .990** | .981** | .977** |
| 55 zad        |        |        |        | 1      | .992** | .982** | .979** |
| 50 zad        |        |        |        |        | 1      | .990** | .986** |
| 45 zad        |        |        |        |        |        | 1      | .994** |
| 40 zad        |        |        |        |        |        |        | 1      |

\*\* statistički značajno na razini 0.01

## 5. RASPRAVA

Glavni cilj ovoga rada bio je provesti kompletnu analizu zadataka novih formi Problemnog testa, forme Pb-94 i forme Pb-95, koja još nigdje nije objavljena. Osim provjere funkcioniranja zadataka, nastojalo se predložiti skraćenu verziju Problemnog testa te provjeriti promjenu u deskriptivnoj statistici dviju verzija tog testa. Također, provedena je i faktorska analiza s ciljem provjere konstruktne valjanosti testova.

Jedan od većih nedostataka ovoga istraživanja je nereprezentativnost uzorka. Takav uzorak, naravno, ne može predstavljati populaciju te nam dobiveni podaci jako malo mogu reći o ispitivanoj sposobnosti u populaciji. Dakle, generalizacija podataka je ograničena, no usporedbu je dvaju testova opravdano provoditi zbog sličnosti uzorka, tj. uzorci su izjednačeni po određenim parametrima. Jedini problem je različita veličina uzorka, no u oba slučaja broj sudionika je dosta velik.

Temeljita deskriptivna statistika dviju novih verzija Problemnog testa već je provedena i objavljena (vidi Kulenović i Žebec, 1998), tako da tome nije posvećena prevelika pažnja u ovome radu. Samo radi lakšeg snalaženja u daljnjem čitanju prikazani su osnovni parametri testova. Već prije provedeni testovi statističke značajnosti govore o paralelnom funkcioniranju parametara u oba testa. Također su oba visoko pouzdana te statistički značajno povezana s korištena tri vanjska kriterija (ispiti znanja iz psihologije, biologije i matematike). Prosječna interkorelacija zadataka u oba testa je slična i niska te su čak neki zadaci u statistički značajnoj negativnoj korelaciji. To isto može biti posljedica korištenja binarnih zadataka koji inače imaju niske prosječne interkorelacije.

Veći je interes ovoga rada analiza testova na razini zadataka. Vizualni pregled indeksa lakoće te koeficijenata diskriminativne valjanosti zadataka ocrta djelomičnu nejednakost zadataka u dvama testovima. Takva razlika istog zadatka u oba testa (koji su intencionalno paralelni) može govoriti da nešto ne valja sa sadržajem toga zadatka u jednom od testova te ih je potrebno sadržajno provjeriti i eventualno ispraviti. Uzroci tome mogu biti razni, kao npr. atraktivnost ponuđenih alternativnih opcija u zadatku ili nerazumljivost samog zadatka. Dakle, potrebna je detaljnija sadržajna analiza zadataka koji se bitno razlikuju u navedenim parametrima. Takva različitost intencionalno paralelnih zadataka može biti jedan od izvora problema pri skraćivanju testova.

Težina zadataka se povećava prema kraju testa te je to slučaj kod oba testa. Postoji nekoliko argumenata u prilog i nekoliko protiv takvog rasporeda zadataka. Argument u prilog bio bi da ispitanik dobije osjećaj uspjeha na početku ili da se točno provjeri do koje težine je ispitanik u stanju rješavati zadatke. Argumenti protiv bili bi da se sudionik “istroši” na laganim zadacima i potroši vrijeme na njih te, ako i stigne do težih, nema ih snage i vremena rješavati. Zadaci koji dobro funkcioniraju u obama testovima nalaze se uglavnom u drugom dijelu što znači da će se izbacivanjem zadataka povećavati težina testa. Isto tako se kod skraćanja testa javlja problem prvih nekoliko zadataka, kojima je indeks lakoće maksimalan, a koeficijent diskriminativne valjanosti minimalan (nekima čak i negativan). Ti zadaci imaju svoju ulogu na početku testa, kao što je već navedeno, no oni djelomično narušavaju pouzdanost cijeloga testa. Jedan je od prijedloga da se ti zadaci izbace iz obrade (ne budu) no zadrže u testu.

Koeficijenti diskriminativne valjanosti su različitih veličina, a u oba testa zajedno oni variraju od -0.0482 do 0.4979. Negativan koeficijent diskriminativne valjanosti, iako je premali da bude statistički značajan, ukazuje na negativnu povezanost zadatka s ukupnim uratkom. Takve zadatke također treba ili popraviti ili izbaciti.

Primijenjena je faktorska analiza na dvjema novim verzijama Problemnog testa, između ostalog jer nigdje ne postoje podaci o njihovoj konstruktnoj valjanosti. Nakon provedenih testova pogodnosti korelacijske matrice za provedbu faktorske analize te potvrdnih rezultata, pristupilo se postupku analize. Faktorska analiza na Problemnom testu A-forme (vidi Krković i Kolesarić, 1973) ukazala je na jedan dominantni faktor koji objašnjava najveći dio varijance te je nazvan “osjetljivost za probleme”, te na još dva manja faktora s osjetno manjim udjelima, koji su nazvani “sposobnost misaonog manipuliranja verbalnim simbolima” i “sposobnost edukcije konektiva koji povezuje manipuliranja verbalnim simbolima”. U ovome radu, faktorska analiza pod modelom glavnih komponenata na dvjema novim formama Problemnog testa uz Guttman-Kaiserov kriterij kod oba testa rezultira većim brojem faktora (za Pb-94 dobiveno je 26 faktora, a za Pb-95 27 faktora). No količina objašnjene varijance i veličina karakterističnih korjenova ukazuje na postojanje jednog značajnog faktora. Pretpostavljamo da se radi o glavnom faktoru kojeg se može nazvati “osjetljivost za probleme”. Zadaci koji su u oba testa najviše povezani s glavnim faktorom ujedno su i oni najdiskriminativniji, tako da se pri skraćanju testa može

oslanjati i na zasićenje glavnim faktorom. Paralelno s tim, promatrajući zasićenja svakog zadatka opet se uočava razlika između istog zadatka u dvama testovima.

Izrada dvije skraćene verzije korištenih testova je nešto kompliciranija jer se tijekom odabira najmanje i najviše funkcionalnih zadataka nailazi na mnoge probleme. Neki od njih su već navedeni, no ovo je mjesto da ih zajedno ponovno spomenemo. Naime, određeni dio zadataka različito funkcionira u dva testa, tako da njegovim izbacivanjem radimo dobro jednome testu, a loše drugome. Osim toga, zadaci koji dobro funkcioniraju u oba testa uglavnom se nalaze u drugom dijelu testa koji je nešto teži od prvoga dijela. Stoga je neminovno da se izbacivanjem nevaljanih zadataka povećava težina testa. Kao najbolji primjer za navedeno, bilo bi prvih nekoliko zadataka koji dosta loše funkcioniraju. Ali budući da su oni ujedno i najlakši, njihovim izbacivanjem uklanja se psihičko olakšanje ispitanika pri početku rješavanja testa.

Pokušali smo skratiti testove, tako da izbacimo iste zadatke u oba testa, a da se osnovne karakteristike samih testova bitno ne promijene. Također se nastojalo da testovi nakon skraćivanja ostanu i dalje paralelni. Kao što je prikazano u obradi rezultata, skraćenje je napravljeno u nekoliko koraka, tako da se postupno izbacivalo po pet zadataka te su se provjeravale promjene u osnovnim parametrima testova. Izbacivalo se najprije zadatke koji podjednako loše funkcioniraju u oba testa, zatim one koji loše funkcioniraju barem u jednome od njih. No treba napomenuti da su ponuđena i provedena skraćjenja samo jedan od mnogih načina skraćjenja korištenih testova te vrlo vjerojatno nisu optimalna.

Aritmetička sredina testa se smanjuje kao i prosječni indeks lakoće. To se moglo očekivati, no u tome određenu ulogu ima i činjenica da je većina izbačenih zadataka bila ona s većim indeksom lakoće, što se vidi i iz smanjenja prosječnog indeksa lakoće. Iako je primijećen navedeni pad u prosječnom indeksu lakoće, on se u biti s viših vrijednosti približio srednjim vrijednostima, što znači da su zadaci postali, što je i poželjnije, srednje teški. Promjena u pouzdanosti testova je ponešto različita kod dvaju testova. Iako te promjene nisu velikih razmjera, to bi mogao biti još jedan primjer nejednakog funkcioniranja testova na razini zadataka, jer većina izbačenih zadataka su vjerojatno oni koji loše funkcioniraju u Pb-94, ali dosta dobro u Pb-95. Ipak, kao što je već navedeno, korisno bi bilo provesti sadržajnu analizu zadataka, da se provjeri zašto određeni zadaci loše funkcioniraju u jednom testu, a izvrsno u drugome.

Interkorelacije zadataka sa skraćanjem testova rastu u oba testa, što upućuje na povećanje dijagnostičke valjanosti testova te njihovo faktorsko čišćenje. Pošto je Cronbachov  $\alpha$  izbacivanjem zadataka na početku rastao, pretpostavlja se da su većina izbačenih zadataka bili u niskim korelacijama s drugim zadacima, pošto se Cronbachov koeficijent pouzdanosti računa na temelju međusobnih interkorelacija zadataka.

Smanjuju se također raspon rezultata, standardna devijacija te BOR, a minimum i maksimum rezultat pokazuju da su se rezultati pomaknuli prema nižim vrijednostima. No Fergusonov koeficijent osjetljivosti, koji u svom računu neutralizira utjecaj smanjenja broja zadataka, ne ukazuje na promjene u osjetljivosti. S obzirom na navedeno, može se reći da se osjetljivost testa izbacivanjem zadataka znatnije ne mijenja. Primjećuje se određeni pad, no to je uglavnom zbog same činjenice da se smanjuje broj zadataka. Unatoč tome da se distribucija rezultata pomiče s viših k srednjim vrijednostima, uspoređivanjem s normalnom distribucijom nisu pronađena statistički značajna odstupanja ni u jednoj verziji oba testa.

Prognostička valjanost testova se izbacivanjem zadataka znatnije ne mijenja stoga jer se korelacije s tri vanjska kriterija uglavnom kreću oko istih vrijednosti. Konstruktna valjanost se također skraćanjem testova ne mijenja. Jedino što se događa jest da se ukupan broj faktora s karakterističnim korijenom većim od 1 smanjuje te ponekad drugi faktor dobiva više na vrijednosti, ali još uvijek ne dovoljno da bi se moglo smisleno i jasno objasniti njegovo postojanje.

Ponovljenom provjerom paralelnosti svih skraćenih verzija testova Pb-94 i Pb-95 s obzirom na određene parametre, dobiveno je da testovi zadržavaju svoju osobitost paralelnosti na svim verzijama skraćanja. Koju od skraćenih verzija bi se prihvatilo ili koristilo, ovisi o mnogo stvari, prije svega o cilju upotrebe, vremenskom ograničenju ili slično. U bilo kojem slučaju, s obzirom na dobivene rezultate, korištenjem kraćih verzija neće se izgubiti na kvaliteti dobivenih podataka.

Međusobne korelacije svih verzija svakog od testa govore o njihovoj visokoj povezanosti. Iako se radi o spurioznim korelacijama, visoki koeficijenti korelacije znače da su izbačeni zadaci koji nisko koreliraju s ukupnim rezultatom. Ono što pruža najviše informacija jest korelacije testa sa svim zadacima i s najmanjim brojem zadataka, koje su također dosta visoke. Dakle, napravljen je dobar odabir zadataka koji će se izbaciti.

Za daljnja istraživanja, predložila bih kompletnu standardizaciju cijelih verzija testova Pb-94 te Pb-95 s obzirom na spol, dob, mjesto stanovanja, struku i ostale demografske varijable. Na tim sakupljenim podacima potrebna je ponovna provjera funkcioniranja kraćih verzija. Isto tako, trebalo bi se primijeniti testove Pb-94 te Pb-95 na zavisnom uzorku radi dobivanja boljih dokaza o njihovoj paralelnosti, kao i izračunavanje njihove međusobne korelacije. Preporučuje se korištenje istih vanjskih kriterija. Iako je to već teže izvedivo, bilo bi korisno na jednom zavisnom uzorku primijeniti testove Pb-94, Pb-95 no i Problemni test A-forme, radi provjere paralelnosti novih testova sa starim. Na tim rezultatima bi se trebala provjeriti i faktorska analiza navedenih testova, jer faktorska analiza Problemnog testa A-forme govori o 3 veća faktora (Krković i Kolesarić, 1973), a u ovome radu dobiven je svega jedan.



## 6. ZAKLJUČAK

Analiza zadataka pokazala je da se zadaci u određenoj mjeri razlikuju između dvije nove forme Problemnog testa. Indeks težine u oba testa zajedno varira od 0.0371 do 0.9971. Smanjuje se prema kraju testa. Koeficijenti diskriminativne valjanosti u oba testa nalaze se u rasponu između -0.0482 i 0.4979. Interkorelacija zadataka kreću se od -0.143 do 0.466.

Faktorskom analizom, uz Guttman-Kaiserov kriterij, dobiva se u testu Pb-94 26 faktora, a u testu Pb-95 27 faktora. No, veličine njihovih karakterističnih korjenova ukazuju na jedan glavni faktor te ostale koji se ne razlikuju puno po svojim navedenim vrijednostima. Dakle, s obzirom na količinu objašnjene ukupne varijance i s obzirom na scree-plot faktora, moglo bi se reći da se ističe samo jedan faktor.

Postepeno izbacivanje zadataka iz obje forme testa rezultiralo je povećanjem težine testa, tj. distribucija se pomiče prema srednjim vrijednostima. Promjene u Cronbachovom  $\alpha$  kod oba testa nisu velike. Prosječna interkorelacija zadataka s izbacivanjem zadataka konstantno raste u oba testa, a korelacije oba testa s tri vanjska kriterija se uz male pomake uglavnom održavaju stalnima. Faktorska struktura se promjenom broja zadataka uglavnom ne mijenja. Raspon rezultata, standardna devijacija i BOR testova se smanjuju, no Fergusonov koeficijent se ne mijenja skraćanjem testa.

Izbacivanjem čak 30 zadataka, osnovne metrijske karakteristike testova se znatnije ne mijenjaju. Isto tako, skraćanjem testova dvije nove forme Problemnog testa ipak ostaju paralelne u svim verzijama s različitim brojem zadataka.

Preporučuje se provesti standardizaciju na reprezentativnom uzorku, kao i ponovnu provjeru dobivenih rezultata u ovome radu također na reprezentativnijem uzorku.

## 7. PRILOG

**Tablica 13.** Promjena u Cronbachovom  $\alpha$  ako se pojedini zadatak izbací iz testa.

| zad                | Cronbach $\alpha$ Pb94<br>ako se izbací zadatak | Cronbach $\alpha$ Pb95<br>ako se izbací zadatak |
|--------------------|---|---|
| <b>cijeli test</b> | <b>0.8655</b>                                   | <b>0.8659</b>                                   |
| 1                  | .8657   | .8662   |
| 2                  | .8652   | .8654   |
| 3                  | .8650   | .8662   |
| 4                  | .8629   | .8639   |
| 5                  | .8638   | .8656   |
| 6                  | .8653   | .8655   |
| 7                  | .8638   | .8625   |
| 8                  | .8652   | .8649   |
| 9                  | .8619   | .8640   |
| 10                 | .8691   | .8658   |
| 11                 | .8650   | .8652   |
| 12                 | .8652   | .8652   |
| 13                 | .8638   | .8644   |
| 14                 | .8612   | .8642   |
| 15                 | .8649   | .8662   |
| 16                 | .8616   | .8637   |
| 17                 | .8654   | .8662   |
| 18                 | .8651   | .8634   |
| 19                 | .8653   | .8658   |
| 20                 | .8653   | .8651   |
| 21                 | .8640   | .8635   |
| 22                 | .8648   | .8643   |
| 23                 | .8650   | .8641   |
| 24                 | .8611   | .8618   |
| 25                 | .8617   | .8648   |
| 26                 | .8639   | .8642   |
| 27                 | .8636   | .8622   |
| 28                 | .8642   | .8630   |
| 29                 | .8674   | .8657   |
| 30                 | .8643   | .8649   |
| 31                 | .8630   | .8662   |
| 32                 | .8648   | .8652   |
| 33                 | .8642   | .8656   |
| 34                 | .8647   | .8638   |
| 35                 | .8660   | .8644   |
| 36                 | .8636   | .8661   |
| 37                 | .8602   | .8615   |
| 38                 | .8630   | .8637   |
| 39                 | .8639   | .8643   |
| 40                 | .8645   | .8658   |
| 41                 | .8642   | .8631   |
| 42                 | .8633   | .8630   |
| 43                 | .8662   | .8628   |
| 44                 | .8601   | .8611   |
| 45                 | .8657   | .8642   |
| 46                 | .8659   | .8647   |

|    |       |       |
|----|-------|-------|
| 47 | .8640 | .8645 |
| 48 | .8633 | .8617 |
| 49 | .8647 | .8659 |
| 50 | .8623 | .8651 |
| 51 | .8609 | .8615 |
| 52 | .8640 | .8680 |
| 53 | .8637 | .8641 |
| 54 | .8607 | .8630 |
| 55 | .8625 | .8642 |
| 56 | .8616 | .8634 |
| 57 | .8641 | .8638 |
| 58 | .8613 | .8625 |
| 59 | .8619 | .8645 |
| 60 | .8619 | .8639 |
| 61 | .8619 | .8647 |
| 62 | .8642 | .8635 |
| 63 | .8647 | .8634 |
| 64 | .8643 | .8638 |
| 65 | .8647 | .8641 |
| 66 | .8647 | .8655 |
| 67 | .8623 | .8622 |
| 68 | .8643 | .8635 |
| 69 | .8609 | .8612 |
| 70 | .8628 | .8633 |

## 8. LITERATURA

Grgin, T. (1989). VII Dani psihologije u Zadru. 5, 70-73

Guilford, J.P. (1971). *The nature of human intelligence*. London: McGraw-Hill.

Krković, A. & Kolesarić, V. (1970). Prikaz novog testa za ispitivanje osjetljivosti za probleme. *Revija za psihologiju*, 1(2), 73-78.

Krković, A. & Kolesarić, V. (1973). Prilog određivanju simptomatske valjanosti Problemnog testa i M-serije. Zbornik radova "*Dani Ramira Bujasa 1970. i 1972.*". Zagreb: Društvo psihologa SR Hrvatske, 175-178.

Kulenović, A. & Krizmanić, M. (1991). Zoran Bujas' work on the construction of psychological measurement instruments. *Revija za psihologiju* 21(1), 25-28.

Kulenović, A. & Žebec, M.S. (1998). Empirijska provjera ekvivalentnosti dvije nove forme Problemnog testa. *Suvremena psihologija*, 1(1-2), 31-42.

Petz, B. (1992). *Psihologijski rječnik*. Zagreb: Prosvjeta

Sobota, I. (1998). *Problemni test (A) - priručnik*. Hrvatski zavod za zapošljavanje, Zagreb

Zarevski, P. (2000). *Struktura i priroda inteligencije*. Jastrebarsko: Naklada Slap

Žebec, M.S. (1997). *Empirijsko ispitivanje paralelnosti dvije nove forme Problemnog testa*. Diplomski rad. Filozofski fakultet, Zagreb.